

0590
0423

#6



OIPE

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/905,176

DATE: 04/23/2002

TIME: 15:30:15

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04232002\I905176.raw

3 <110> APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY
 4 Debe, Derek A.
 6 <120> TITLE OF INVENTION: METHOD FOR DETERMINING THREE-DIMENSIONAL PROTEIN STRUCTURE
 FROM PRIMARY
 7 PROTEIN SEQUENCE
 9 <130> FILE REFERENCE: 265/297
 11 <140> CURRENT APPLICATION NUMBER: US 09/905,176
 C--> 12 <141> CURRENT FILING DATE: 2002-04-05
 14 <150> PRIOR APPLICATION NUMBER: US 60/218,016
 15 <151> PRIOR FILING DATE: 2000-07-12
 17 <160> NUMBER OF SEQ ID NOS: 26
 19 <170> SOFTWARE: PatentIn version 3.1
 21 <210> SEQ ID NO: 1
 22 <211> LENGTH: 53
 23 <212> TYPE: PRT
 24 <213> ORGANISM: Artificial Sequence
 26 <220> FEATURE:
 27 <223> OTHER INFORMATION: Genus/species, Unknown
 29 <400> SEQUENCE: 1
 31 Leu Val Ala Phe Ala Asp Phe Gly Ser Val Thr Phe Thr Asn Ala Glu
 32 1 5 10 15
 35 Ala Thr Ser Gly Gly Ser Thr Val Gly Pro Ser Asp Ala Thr Val Met
 36 20 25 30
 39 Asp Ile Glu Gln Asp Gly Ser Val Leu Thr Glu Thr Ser Val Ser Gly
 40 35 40 45
 43 Asp Ser Val Thr Val
 44 50
 47 <210> SEQ ID NO: 2
 48 <211> LENGTH: 53
 49 <212> TYPE: PRT
 50 <213> ORGANISM: Artificial Sequence
 52 <220> FEATURE:
 53 <223> OTHER INFORMATION: Genus/species, Unknown
 55 <400> SEQUENCE: 2
 57 Leu Val Ala Phe Ala Asp Phe Gly Ser Val Thr Phe Thr Asn Ala Glu
 58 1 5 10 15
 61 Ala Thr Ser Gly Gly Ser Thr Val Gly Pro Ser Asp Ala Thr Val Met
 62 20 25 30
 65 Asp Ile Glu Gln Asp Gly Ser Val Leu Thr Glu Thr Ser Val Ser Gly
 66 35 40 45
 69 Asp Ser Val Thr Val
 70 50
 73 <210> SEQ ID NO: 3
 74 <211> LENGTH: 53

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/905,176

DATE: 04/23/2002

TIME: 15:30:15

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04232002\I905176.raw

```

75 <212> TYPE: PRT
76 <213> ORGANISM: Artificial Sequence
78 <220> FEATURE:
79 <223> OTHER INFORMATION: Genus/species, Unknown
81 <400> SEQUENCE: 3
83 Leu Val Pro Phe Ala Asn Phe Gly Thr Val Thr Phe Thr Gly Ala Glu
84 1 5 10 15
87 Ala Thr Thr Ser Ser Gly Thr Val Thr Ala Ala Asp Ala Thr Leu Ile
88 20 25 30
91 Asp Ile Glu Gln Asn Gly Glu Val Leu Thr Ser Val Thr Val Ser Gly
92 35 40 45
95 Ser Thr Val Thr Val
96 50
99 <210> SEQ ID NO: 4
100 <211> LENGTH: 52
101 <212> TYPE: PRT
102 <213> ORGANISM: Artificial Sequence
104 <220> FEATURE:
105 <223> OTHER INFORMATION: Genus/species, Unknown
107 <400> SEQUENCE: 4
109 Leu Val Gln Phe Ala Asn Phe Gly Thr Val Thr Phe Thr Gly Ala Ser
110 1 5 10 15
113 Ala Thr Gln Asn Gly Glu Ser Val Gly Val Thr Gly Ala Gln Ile Ile
114 20 25 30
117 Asp Leu Gln Gln Asn Ser Val Leu Thr Ser Val Ser Thr Ser Ser Asn
118 35 40 45
121 Ser Val Thr Val
122 50
125 <210> SEQ ID NO: 5
126 <211> LENGTH: 47
127 <212> TYPE: PRT
128 <213> ORGANISM: Artificial Sequence
130 <220> FEATURE:
131 <223> OTHER INFORMATION: Genus/species, Unknown
133 <400> SEQUENCE: 5
135 Leu Val Asn Phe Ala Asp Phe Asp Thr Val Thr Phe Lys Asp Cys Ser
136 1 5 10 15
139 Pro Ser Val Ser Gly Ser Thr Ile Val Asp Ile Arg Gln Ser Leu Glu
140 20 25 30
143 Val Leu Thr Glu Cys Ser Thr Thr Gly Thr Thr Thr Val Thr Cys
144 35 40 45
147 <210> SEQ ID NO: 6
148 <211> LENGTH: 54
149 <212> TYPE: PRT
150 <213> ORGANISM: Artificial Sequence
152 <220> FEATURE:
153 <223> OTHER INFORMATION: Genus/species, Unknown
155 <400> SEQUENCE: 6
157 Phe Val Pro Phe Ala Ser Phe Ser Pro Ala Val Glu Phe Thr Asp Cys

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/905,176

DATE: 04/23/2002

TIME: 15:30:15

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04232002\I905176.raw

```

158 1          5          10          15
161 Ser Val Thr Ser Asp Gly Glu Ser Val Ser Leu Asp Asp Ala Gln Ile
162          20          25          30
165 Thr Gln Val Ile Ile Asn Asn Gln Asp Val Thr Asp Cys Ser Val Ser
166          35          40          45
169 Gly Thr Thr Val Ser Cys
170          50
173 <210> SEQ ID NO: 7
174 <211> LENGTH: 54
175 <212> TYPE: PRT
176 <213> ORGANISM: Artificial Sequence
178 <220> FEATURE:
179 <223> OTHER INFORMATION: Genus/species, Unknown
181 <400> SEQUENCE: 7
183 Phe Val Pro Phe Ala Ser Phe Ser Pro Ala Val Glu Phe Thr Asp Cys
184 1          5          10          15
187 Ser Val Thr Ser Asp Gly Glu Ser Val Ser Leu Asp Asp Ala Gln Ile
188          20          25          30
191 Thr Gln Val Ile Ile Asn Asn Gln Asp Val Thr Asp Cys Ser Val Ser
192          35          40          45
195 Gly Thr Thr Val Ser Cys
196          50
199 <210> SEQ ID NO: 8
200 <211> LENGTH: 54
201 <212> TYPE: PRT
202 <213> ORGANISM: Pseudomonas aeruginosa
204 <400> SEQUENCE: 8
206 Phe Val Pro Phe Ala Ser Phe Ser Pro Ala Val Glu Phe Thr Asp Cys
207 1          5          10          15
210 Ser Val Thr Ser Asp Gly Glu Ser Val Ser Leu Asp Asp Ala Gln Ile
211          20          25          30
214 Thr Gln Val Ile Ile Asn Asn Gln Asp Val Thr Asp Cys Ser Val Ser
215          35          40          45
218 Gly Thr Thr Val Ser Cys
219          50
222 <210> SEQ ID NO: 9
223 <211> LENGTH: 326
224 <212> TYPE: PRT
225 <213> ORGANISM: Saccharomyces cerevisiae
227 <400> SEQUENCE: 9
229 Tyr Pro Tyr Thr Arg Leu Arg Arg Asn Arg Arg Asp Asp Phe Ser Arg
230 1          5          10          15
233 Arg Leu Val Arg Glu Asn Val Leu Thr Val Asp Asp Leu Ile Leu Pro
234          20          25          30
237 Val Phe Val Leu Asp Gly Val Asn Gln Arg Glu Ser Ile Pro Ser Met
238          35          40          45
241 Pro Gly Val Glu Arg Leu Ser Ile Asp Gln Leu Leu Ile Glu Ala Glu
242          50          55          60
245 Glu Trp Val Ala Leu Gly Ile Pro Ala Leu Ala Leu Phe Pro Val Thr

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/905,176

DATE: 04/23/2002

TIME: 15:30:15

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04232002\I905176.raw

```

246 65          70          75          80
249 Pro Val Glu Lys Lys Ser Leu Asp Ala Ala Glu Ala Tyr Asn Pro Glu
250          85          90          95
253 Gly Ile Ala Gln Arg Ala Thr Arg Ala Leu Arg Glu Arg Phe Pro Glu
254          100          105          110
257 Leu Gly Ile Ile Thr Asp Val Ala Leu Asp Pro Phe Thr Thr His Gly
258          115          120          125
261 Gln Asp Gly Ile Leu Asp Asp Asp Gly Tyr Val Leu Asn Asp Val Ser
262          130          135          140
265 Ile Asp Val Leu Val Arg Gln Ala Leu Ser His Ala Glu Ala Gly Ala
266 145          150          155          160
269 Gln Val Val Ala Pro Ser Asp Met Met Asp Gly Arg Ile Gly Ala Ile
270          165          170          175
273 Arg Glu Ala Leu Glu Ser Ala Gly His Thr Asn Val Arg Ile Met Ala
274          180          185          190
277 Tyr Ser Ala Lys Tyr Ala Ser Ala Tyr Tyr Gly Pro Phe Arg Asp Ala
278          195          200          205
281 Val Gly Ser Ala Ser Asn Leu Gly Lys Gly Asn Lys Ala Thr Tyr Gln
282          210          215          220
285 Met Asp Pro Ala Asn Ser Asp Glu Ala Leu His Glu Val Ala Ala Asp
286 225          230          235          240
289 Leu Ala Glu Gly Ala Asp Met Val Met Val Lys Pro Gly Met Pro Tyr
290          245          250          255
293 Leu Asp Ile Val Arg Arg Val Lys Asp Glu Phe Arg Ala Pro Thr Phe
294          260          265          270
297 Val Tyr Gln Val Ser Gly Glu Tyr Ala Met His Met Gly Ala Ile Gln
298          275          280          285
301 Asn Gly Trp Leu Ala Glu Ser Val Ile Leu Glu Ser Leu Thr Ala Phe
302          290          295          300
305 Lys Arg Ala Gly Ala Asp Gly Ile Leu Thr Tyr Phe Ala Lys Gln Ala
306 305          310          315          320
309 Ala Glu Gln Leu Arg Arg
310          325
313 <210> SEQ ID NO: 10
314 <211> LENGTH: 328
315 <212> TYPE: PRT
316 <213> ORGANISM: Saccharomyces cerevisiae
318 <400> SEQUENCE: 10
320 Glu Ile Ser Ser Val Leu Ala Gly Gly Tyr Asn His Pro Leu Leu Arg
321 1          5          10          15
324 Gln Trp Gln Ser Glu Arg Gln Leu Thr Lys Asn Met Leu Ile Phe Pro
325          20          25          30
328 Leu Phe Ile Ser Asp Asn Pro Asp Asp Phe Thr Glu Ile Asp Ser Leu
329          35          40          45
332 Pro Asn Ile Asn Arg Ile Gly Val Asn Arg Leu Lys Asp Tyr Leu Lys
333          50          55          60
336 Pro Leu Val Ala Lys Gly Leu Arg Ser Val Ile Leu Phe Gly Val Pro
337 65          70          75          80
340 Leu Ile Pro Gly Thr Lys Asp Pro Val Gly Thr Ala Ala Asp Asp Pro

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/905,176

DATE: 04/23/2002

TIME: 15:30:15

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04232002\I905176.raw

```

341          85          90          95
344 Ala Gly Pro Val Ile Gln Gly Ile Lys Phe Ile Arg Glu Tyr Phe Pro
345          100          105          110
348 Glu Leu Tyr Ile Ile Cys Asp Val Cys Leu Cys Glu Tyr Thr Ser His
349          115          120          125
352 Gly His Cys Gly Val Leu Tyr Asp Asp Gly Thr Ile Asn Arg Glu Arg
353          130          135          140
356 Ser Val Ser Arg Leu Ala Val Ala Val Asn Tyr Ala Lys Ala Gly
357 145          150          155          160
360 Ala His Cys Val Ala Pro Ser Asp Met Ile Asp Gly Arg Ile Arg Asp
361          165          170          175
364 Ile Lys Arg Gly Leu Ile Asn Ala Asn Leu Ala His Lys Thr Phe Val
365          180          185          190
368 Leu Ser Tyr Ala Ala Lys Phe Ser Gly Asn Leu Tyr Gly Pro Phe Arg
369          195          200          205
372 Asp Ala Ala Cys Ser Ala Pro Ser Asn Gly Asp Arg Lys Cys Tyr Gln
373          210          215          220
376 Leu Pro Pro Ala Gly Arg Gly Leu Ala Arg Arg Ala Leu Glu Arg Asp
377 225          230          235          240
380 Met Ser Glu Gly Ala Asp Gly Ile Ile Val Lys Pro Ser Thr Phe Tyr
381          245          250          255
384 Leu Asp Ile Met Arg Asp Ala Ser Glu Ile Cys Lys Asp Leu Pro Ile
385          260          265          270
388 Cys Ala Tyr His Val Ser Gly Glu Tyr Ala Met Leu His Ala Ala Ala
389          275          280          285
392 Glu Lys Gly Val Val Asp Leu Lys Thr Ile Ala Phe Glu Ser His Gln
393          290          295          300
396 Gly Phe Leu Arg Ala Gly Ala Arg Leu Ile Ile Thr Tyr Leu Ala Pro
397 305          310          315          320
400 Glu Phe Leu Asp Trp Leu Asp Glu
401          325
404 <210> SEQ ID NO: 11
405 <211> LENGTH: 215
406 <212> TYPE: PRT
407 <213> ORGANISM: Homo sapiens
409 <400> SEQUENCE: 11
411 Ala Gly His Asp Asn Thr Lys Pro Asp Thr Ser Ser Ser Leu Leu Thr
412 1          5          10          15
415 Ser Leu Asn Gln Leu Gly Glu Arg Gln Leu Leu Ser Val Val Lys Trp
416          20          25          30
419 Ser Lys Ser Leu Pro Gly Phe Arg Asn Leu His Ile Asp Asp Gln Ile
420          35          40          45
423 Thr Leu Ile Gln Tyr Ser Trp Met Ser Leu Met Val Phe Gly Leu Gly
424          50          55          60
427 Trp Arg Ser Tyr Lys His Val Ser Gly Gln Met Leu Tyr Phe Ala Pro
428 65          70          75          80
431 Asp Leu Ile Leu Asn Glu Gln Arg Met Lys Glu Ser Ser Phe Tyr Ser
432          85          90          95
435 Leu Cys Leu Thr Met Trp Gln Ile Pro Gln Glu Phe Val Lys Leu Gln

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/905,176

DATE: 04/23/2002

TIME: 15:30:16

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04232002\I905176.raw

L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date